AD-A115 347 STANFORD UNIV CA DEPT OF STATISTICS F/G 12/1

A NOTE ON NON-LINEAR FUNCTIONS OF LINEAR COMBINATIONS.(U)

FEB 82 P DIACONIS, M SHASHAHANI N00014-76-C-0475

UNCLASSIFIED TR-313 NL

1 OF 1

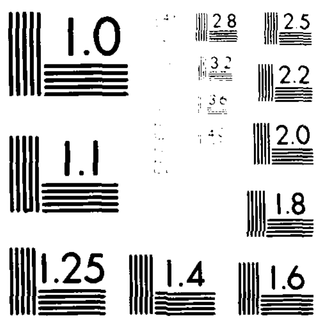
AD A
115347

END

DATE FILMED

07-82

DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

A NOTE ON NON-LINEAR FUNCTIONS OF LINEAR COMBINATIONS

By

Persi Diaconis and Mehrdad Shashahani

TECHNICAL REPORT NO. 313

February 2, 1982

Prepared Under Contract
N00014-76-C-0475 (NR-042-267)
For the Office of Naval Research

Herbert Solomon, Project Director

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

A NOTE ON NON-LINEAR FUNCTIONS OF LINEAR COMBINATIONS

Persi Diaconis

Harvard University and Stanford University

Mehrdad Shahshahani

University of Virginia and Stanford University

ABSTRACT

Projection pursuit algorithms approximate a function of p variables by a sum of non-linear functions of linear combinations

$$(1) \qquad f(x_1, \ldots, x_p) \doteq \sum_{i=1}^{n} g_i(a_{i1}x_1 + \ldots + a_{ip}x_p) \; .$$

We develop some approximation theory, give a necessary and sufficient condition for equality in (1), and discuss non-uniqueness of the representation.

1. Introduction and Statement of Main Results

We present some mathematical analysis for a class of curve fitting algorithms labeled "projection pursuit" algorithms by Friedman and Stuetzle (1981 a, b). These algorithms approximate a general function of p variables by a sum of non-linear functions of projections:

$$f(x_1, \ldots, x_p) \doteq \sum_{i=1}^{n} g_i(a_{i1}x_1 + \ldots + a_{ip}x_p) . \tag{1.1}$$

In (1.1), f is a given function and univariate, non-linear functions g_i and linear combinations $a_{i1}x_1 + \ldots + a_{ip}x_p$ are sought so that a reasonable approximation is attained. Such approximation is computationally feasible and performs well in examples of non-linear regression with noisy data, high dimensional density estimation, and multidimensional splines. In addition to the articles of Friedman and Stuetzle cited above see Friedman and Tukey (1974), Friedman, Gross and Stuetzle (1981) for examples and computational details. Huber (1981 a, b) begins to connect the algorithms to statistical theory. This note treats the algorithms from the point of view of approximation theory.

It is easy to show that approximation is always possible.

<u>THEOREM 1</u>. Functions of the form $\Sigma\alpha_i e^{\underset{\sim}{a}^i \cdot \underset{\sim}{x}}$, with α_i real, $\underset{\sim}{a}^i$ a vector of nonnegative integers, and $\underset{\sim}{x} = (x_1, \ldots, x_p)$ are dense in the continuous real valued functions on $[0,1]^p$ under the maximum deviation norm.

<u>Proof</u>. The functions $e^{\underset{\sim}{a} \cdot \underset{\sim}{x}}$ separate points of $[0,1]^p$ and are closed under multiplication. Finite linear combinations of such functions form a point separating algebra which is dense because of the Stone-Wierstrass theorem. ∎

<u>THEOREM 2</u>. Functions of the form

$$\Sigma \alpha_i \cos(2\pi \underset{\sim}{a}^i \cdot x) + \beta_i \sin(2\pi b^i \cdot x)$$

are dense in $L^2[0,1]^p$.

<u>Proof</u>. Any function in $L^2[0,1]^p$ can be well approximated by its Fourier expansion. See Volume 2 of Zygmund (1959) and the survey article by Ash (1976) for further details and refinements. ∎

Sometimes equality is possible in (1.1). For example

$$xy = \frac{1}{4}(x+y)^2 - \frac{1}{4}(x-y)^2$$

$$\max(x,y) = \frac{1}{2}|x+y| + \frac{1}{2}|x-y|$$

$$(xy)^2 = \frac{1}{4}(x+y)^4 + \frac{7}{4\cdot 3^3}(x-y)^4 - \frac{1}{2\cdot 3^3}(x+2y)^4 - \frac{2^3}{3^3}(x+\frac{1}{2}y)^4 .$$

In what follows we will focus on conditions for equality in (1.1) as a method of determining examples to test, compare, and evaluate algorithms. Consider first a smooth function of 2 variables of the special form

$$f(x,y) = g(ax+by) .$$

Clearly

$$\left(b\frac{\partial}{\partial x} - a\frac{\partial}{\partial y}\right)f \equiv 0 .$$

If f has the form

$$f(x,y) = \sum_{i=1}^{n} g_i(a_i x + b_i y) \tag{1.2}$$

then the differential operator

$$L = \prod_{i=1}^{n} \left(b_i \frac{\partial}{\partial x} - a_i \frac{\partial}{\partial y} \right) = \sum_{i=0}^{n} c_i \frac{\partial^n}{\partial x^i \partial y^{n-i}}$$

applied to f is identically zero. The next theorem gives a converse.

THEOREM 3. Let $f \in C^k[0,1]^2$. Suppose that for some real numbers $c_0, \ldots, c_n$, the operator $\sum_{i=0}^{n} c_i \frac{\partial^k}{\partial x^i \partial y^{n-i}}$ applied to f is identically zero. If the polynomial $\sum_{i=0}^{n} c_i z^i$ has distinct real zeros then (1.2) holds for some (a_i, b_i). The lines $a_i x + b_i y$ are all distinct.

Theorem 3 is proved in Section 2 which also contains a discussion of techniques for finding directions (a_i, b_i) given f. Some applications of Theorem 3 are contained in the following examples.

APPLICATION 1. The functions e^{xy} and $\sin xy$ cannot be written in the form (1.1) for any finite n. Indeed, the equation $\Sigma c_i \frac{\partial^n}{\partial x^i \partial y^{n-i}} f \equiv 0$ implies $c_i \equiv 0$ and the associated polynomial has complex roots.

APPLICATION 2. Let $f(x,y)$ be a polynomial of degree m. Then

$$f(x,y) = \sum_{i=1}^{m} g_i(a_i x + b_i y)$$

where each g_i is a polynomial of degree at most m. This follows by eliminating manipulations from Theorem 3. Thus, any polynomial in 2 variables can be represented exactly. Since polynomials are dense in $C[0,1]^2$, this gives another proof of denseness of projection pursuit approximations.

APPLICATION 3. Representations of the form (1.1) are not necessarily unique. For example

$$xy = c(ax+by)^2 - c(ax-by)^2$$

for any a and b satisfying $ab \neq 0$, $a^2 + b^2 = 1$ with $c = 1/4ab$.

Writing $a = \cos\theta$, $b = \sin\theta$, any non-coordinate direction can be chosen for the quadratic g_1. The second direction is forced as orthogonal to this. This suggests that substantive interpretation of the linear combinations (a_i, b_i) is difficult. For a more ambitious example, consider the function $(xy)^2$. This is of 4th degree. Use of Theorem 3 as outlined in Section 2, shows that $(xy)^2$ cannot be expressed as a sum of $n = 3$ or fewer terms in (1.1). Four terms of 4th degree suffice:

$$(xy)^2 = \alpha_1(x+b_1y)^4 + \alpha_2(x+b_2y)^4 + \alpha_3(x+b_3y)^4 + \alpha_4(x+b_4y)^4 ,$$

where b_1, b_2, b_3, b_4 are chosen as distinct, and satisfying

$$b_1b_2 + b_1b_3 + b_1b_4 + b_2b_3 + b_2b_4 + b_3b_4 = 0 .$$

Then α_1, α_2, α_3, α_4 are determined by

$$\alpha_i = \frac{1}{6} \frac{\Sigma^* b_j}{\Pi^*(b_j - b_i)}$$

where the sum and product are over $j \neq i$. This clearly defines a three dimensional family of solutions.

APPLICATION 4. Even if the directions (a_i, b_i) are fixed, the representation need not be unique. Suppose that n is the smallest integer such that

$$f(x, y) = \sum_{i=1}^{n} g_i(a_i x + b_i y) .$$

If also

$$f(x, y) = \sum_{i=1}^{n} h_i(a_i x + b_i y) ,$$

then

$$f_i(t) - h_i(t) = p_i(t) , \quad 1 \leq i \leq n$$

with p_i a polynomial of degree at most $n - 1$. The polynomials p_i can be chosen in an arbitrary way subject to the constraint $\Sigma p_i \equiv 0$. In particular, any $n - 1$ of the p_i can be chosen arbitrarily and a final polynomial can be found to satisfy the constraint. These results all follow easily from Theorem 3; indeed the operator $L_i = \prod_{j \neq i} \left(b_j \frac{\partial}{\partial x} - a_j \frac{\partial}{\partial y} \right)$ applied to $f(x, y)$ gives

$$h_i^{(n-1)}(a_i x + b_i y) \prod_{j \neq i} (b_j a_i - a_j b_i) = g_i^{(n-1)}(a_i x + b_i y) \prod_{j \neq i} (b_j a_i - a_j b_i) .$$

The products are non-vanishing because the directions are distinct. It follows that h_i differs from g_i by at most a polynomial of degree $n - 1$, and that an arbitrary polynomial may be added subject to the constraint.

In the special case $n = 2$, Theorem 3 was given by Dotson (1968) who suggests further application to factoring probability densities and separation of variables.

The generalization to dimension greater than two is not as neat. We give the result in 3 dimensions, characterizing functions on $[0, 1]^3$ of the form

$$\sum_{i=1}^{n} g_i(a_{i1} x_1 + a_{i2} x_2 + a_{i3} x_3) . \tag{1.3}$$

Clearly a smooth function $f(x_1, x_2, x_3)$ is of the form $g(x_3)$ if and only if $\frac{\partial}{\partial x_1} f$ and $\frac{\partial}{\partial x_2} f$ vanish identically. It is equivalent to insist that $\left[b_1 \frac{\partial}{\partial x_1} + b_2 \frac{\partial}{\partial x_2} + b_3 \frac{\partial}{\partial x_3}\right] f$ vanishes identically for all (b_1, b_2, b_3) in the plane normal to the x_3 axis (so $b_3 = 0$). The following theorem generalizes these considerations. The generalization to p-dimensions is straightforward.

<u>THEOREM 4</u>. Let Π_i be n distinct planes in $\mathbb{R}^3$. Let $f \in C^n[0, 1]^3$. Then f has the form (1.3) if and only if for all $b_i \in \Pi_i$,

$$\prod_{i=1}^{n} \left\{ b_{i1} \frac{\partial}{\partial x_1} + b_{i2} \frac{\partial}{\partial x_2} + b_{i3} \frac{\partial}{\partial x_3} \right\} f \equiv 0 . \tag{1.4}$$

<u>Remarks</u>. If $c_i, d_i \in \Pi_i$ form a basis, (1.4) holds for all $b_i \in \Pi_i$, $1 \le i \le n$ if and only if it holds for the 2^n cases in which b_i runs over possible basis vectors. The case $n = 2$ in (1.3) is degenerate and may be treated by Theorem 3: for example, a necessary and sufficient condition for $f(x_1, x_2, x_3) = g_1(x_1) + g_2(x_2)$ is $\frac{\partial}{\partial x_3} f$ and $\frac{\partial}{\partial x_1} \frac{\partial}{\partial x_2} f$ vanish identically.

We conclude this introduction by relating the above results to Hilbert's 13th problem. In modern notation Hilbert asked if there are genuine multivariate functions. Of course, $x + y$ is a function of 2 variables but $xy = e^{\log x + \log y}$ is a superposition of univariate functions and $+$. Kolmogorov and Arnold showed that, in this sense, $+$ is the only function of 2 variables. They constructed 5 monotone functions $\phi_i : [0, 1] \to \mathbb{R}$, $\phi_i \in \mathrm{Lip}^1$, with the following remarkable property: for each $f \in C[0, 1]^2$ there is a $g \in C[0, 1]$ such that for all (x, y)

$$f(x, y) = \sum_{i=1}^{n} g\left(\phi_i(x) + \frac{1}{2} \phi_i(y)\right) .$$

Thus ϕ_i are a "universal change of variables" which allows exact equality. A nice discussion of this result and its refinements can be found in Lorentz (1966, 1980) and Vertushkin (1977). While the functions ϕ_i and g are given in a constructive fashion, it does not seem that this result is used to approximate functions in an applied context. This is probably because the functions ϕ_i are fairly "wild". For example, it is known that it is not possible to choose ϕ_i to be C^1 functions, so fixed linear combinations of x and y are ruled out. Indeed, it is known that there is a polynomial $f(x, y)$ for which $f(x, y) = \sum_{i=1}^{n} g_i(a_i x + b_i y)$ is not possible with a_i, b_i chosen independent of f. In the projection pursuit approach to approximation, a_i and b_i are allowed to depend on f and Example 2 shows that now any polynomial can be written in the required form. Example 1 shows that not all functions can be so expressed.

Acknowledgement. We thank Jerry Friedman, Bob Hulquist, and Winnie Li for helpful discussions.

2. Proof and Discussion of Theorems 3 and 4. Let L be the differential operator: $\sum_{i=0}^{n} c_i \frac{\partial^n}{\partial x^i \partial y^{n-i}}$. By hypothesis, the polynomial

$$\sum_{i=0}^{n} c_i x^i y^{n-i} = y^n \sum_{i=0}^{n} c_i \left(\frac{x}{y}\right)^i$$

splits into distinct linear factors. Thus L can be written as $\Pi\left(b_i \frac{\partial}{\partial x} - a_i \frac{\partial}{\partial y}\right)$, with the lines $a_i x + b_i y$ distinct. It must be shown that f can be represented as $\sum_{i=1}^{n} g_i(a_i x + b_i y)$. The proof is by induction on n. For $n = 1$, suppose without real loss that $a_1 \neq 0$. Then $f(x, y) = g(a_1 x + b_1 y)$ with $g(z) = f\left(\frac{z}{a_1}, 0\right)$. One way to show this is to fix (x, y) and define $h(t) = f\left[x + \frac{b_1}{a_1} y - \frac{b_1}{a_1} yt,\ ty\right]$. Then $h(0) = f\left[x + \frac{b_1}{a_1} y,\ 0\right] = g(a_1 x + b_1 y)$; $h(1) = f(x, y)$ and $h'(t) \equiv 0$, for $0 \leq t \leq 1$. The fundamental theorem of calculus gives $h(1) = \int_0^1 h' + h(0)$. Suppose the result is true for operators of degree $\leq n - 1$. To prove it for degree n, write

$$\prod_{i=1}^{n} \left(b_i \frac{\partial}{\partial x} - a_i \frac{\partial}{\partial y}\right) f = \left\{\prod_{i=1}^{n-1} \left(b_i \frac{\partial}{\partial x} - a_i \frac{\partial}{\partial y}\right)\right\} \left(b_n \frac{\partial}{\partial x} - a_n \frac{\partial}{\partial y}\right) f \equiv 0 .$$

By the induction hypothesis, there are functions g_i, $1 \leq i \leq n-1$ satisfying

$$\left(b_n \frac{\partial}{\partial x} - a_n \frac{\partial}{\partial y}\right) f = \sum_{i=1}^{n-1} g_i(a_i x + b_i y) . \tag{2.1}$$

A solution f^* of (2.1) of the form

$$f^*(x, y) = \sum_{i=1}^{n-1} h_i(a_i x + b_i y)$$

is found by choosing $h_i(t) = (b_n a_i - a_n b_i)^{-1} \int_0^t g_i(s)ds$. This is well defined because the lines are distinct. Now $\left\{b_n \frac{\partial}{\partial x} - a_n \frac{\partial}{\partial y}\right\}(f - f^*) \equiv 0$ can be solved explicitly with $f - f^*(x,y) = h_n(a_n x + b_n y)$ by the argument for $n = 1$. It follows that $f = f^* + h_n$ can be written in the required form. ■

Remarks on Explicit Computations. Theorem 3 gives the existence of numbers $c_0, \ldots, c_n$ such that $\Sigma c_j \frac{\partial^n}{\partial x^i \partial y^{n-i}}(f) \equiv 0$. Fixing $n+1$ distinct pairs (x_i, y_i), calculate $\left.\frac{\partial^n}{\partial x^i \partial y^{n-i}}\right|_{(x_i, y_i)}$ and solve the resulting system of equations for c_i. It is feasible to check if the polynomial $c_0 + \ldots + c_n z^n$ has distinct real roots using techniques in Chapter 6 of Henrici (1977). Each stage of the procedure is feasible by a finite algorithm. If the procedure fails at any stage, then equality is impossible. Given feasible $c_0, \ldots, c_n$, it may be possible to find the roots of the associated polynomial. This determines directions (a_i, b_i).

In simple examples there is often enough freedom of choice to make determination of (a_i, b_i) possible. Consider $f(x,y) = xy$ for $n = 2$,

$$\prod_{i=1}^{2}\left(b_i \frac{\partial f}{\partial x} - a_i \frac{\partial f}{\partial y}\right) = b_1 b_2 \frac{\partial^2 f}{\partial x^2} - (b_1 a_2 + b_2 a_1)\frac{\partial^2 f}{\partial x \partial y} + a_1 a_2 \frac{\partial^2 f}{\partial y^2}.$$

Since $\frac{\partial^2 f}{\partial x^2} = \frac{\partial^2 f}{\partial y^2} = 0$, $\frac{\partial^2 f}{\partial x \partial y} = 1$; any distinct choice of a_i and b_i with $b_1 a_2 = -b_2 a_1$ works. Taking $a_1 = b_1 = 1$, $a_2 = -b_2 = 1$, we are led to solve

$$f(x,y) = g_1(x+y) + g_2(x-y).$$

Applying $\frac{\partial}{\partial x} - \frac{\partial}{\partial y}$ to both sides leads to $y - x = 2g_2'(x-y)$; setting $y = 0$; $g_2'(x) = -\frac{x}{2}$, $g_2 = -\frac{x^2}{4} + c_2$. Similarly, $g_1(x) = \frac{x^2}{4} + c_1$ and the result is $xy = \frac{1}{4}(x+y)^2 + c_1 - \frac{1}{4}(x-y)^2 + c_2$ where $c_1 + c_2 = 0$ is forced. In general, if $f = \sum_{i=1}^{n} g_i(a_i x + b_i y)$; $\prod_{j \neq i}\left(b_i \frac{\partial}{\partial x} - a_i \frac{\partial}{\partial y}\right) f = c_i g_i^{(n-1)}(a_i x + b_i y)$, for an explicit c_i. This determines g_i up to an essentially free choice of an $n-1$ degree polynomial.

In the case of a polynomial f, some additional tricks become available. For a multinomial $x^a y^b$ let $a + b = n$; only sums of the form $\sum_{i=1}^{n} \alpha_i (x + \beta_j y)^n$ need be considered. Expanding out and equating coefficients gives

$$\Sigma\alpha_i = 0,\ \Sigma\alpha_i b_i = 0 \ \ldots\ \Sigma\alpha_i \beta_j^a = \frac{1}{\binom{n}{j}} \ \ldots\ \Sigma\alpha_i b_j^n = 0 .$$

This gives $n + 1$ equations in $2n$ unknowns. These are linear in α for β_j given and may be solved explicitly because the matrix is a Vandemonde with a well known inverse. See Goutschi (1963).

Proof of Theorem 4. Condition (1.4) is clearly necessary. For sufficiency, observe first that we may assume that the normals a_i to the planes Π_i span a subspace of dimension 3 or higher. If the dimension of this subspace is two, then the problem reduces to the corresponding problem in $\mathbb{R}^2$ which was solved in Theorem 3. The proof is by induction on n. For $n = 1$, the argument was given in the discussion preceding the theorem. Suppose that the result has been demonstrated for $n - 1$. For $i = 1, \ldots, n$, let b_i and c_i form a basis for Π_i. A generic element of Π_i can be written as $\beta_i b_i + \gamma_i c_i$, for $\beta_i, \gamma_i \in \mathbb{R}$. Write the equation (1.4) as

$$\left(\prod_{i=2}^{n} P_i\right)\left(P_1(f)\right) \equiv 0 . \tag{2.2}$$

By the induction hypothesis,

$$P_1 f = \sum_{i=2}^{n} g_i(a_i \cdot x) . \tag{2.3}$$

We will now find the general solution to (2.3). To begin, note that P_1 may be regarded as a 2-parameter family of differential operators depending linearly on (β_1, γ_1). It follows that the right side of (2.3) must depend linearly on (β_1, γ_1). Write

$$g_i = \beta_1 g_{i1} + \gamma_1 g_{i2} ,$$

and

$$P_1 f = \beta_1 \Sigma g_{i1}(a_i \cdot x) + \gamma_1 \Sigma g_{i2}(a_i \cdot x) .$$

For this equation to have a solution, a necessary integrability condition on g_{i1} and g_{i2} must be satisfied. To see this, write

$$P_1 = \beta_1\left(b_{11} \frac{\partial}{\partial x_1} + b_{12} \frac{\partial}{\partial x_2} + b_{13} \frac{\partial}{\partial x_3}\right) + \gamma_1\left(c_{11} \frac{\partial}{\partial x_1} + c_{12} \frac{\partial}{\partial x_2} + c_{13} \frac{\partial}{\partial x_3}\right) .$$

From (2.3) it must be that the following two equations are satisfied:

$$\sum_{j=1}^{3} b_{1j} \frac{\partial f}{\partial x_j} = \sum_{i=2}^{n} g_{i1}(a_i \cdot x)$$

$$\sum_{j=1}^{3} c_{1j} \frac{\partial f}{\partial x_j} = \sum_{i=2}^{n} g_{i2}(a_i \cdot x) .$$

The necessary condition for integrability is

$$\Sigma c_{1j} \frac{\partial}{\partial x_j} \{ \Sigma g_{i1}(a_i \cdot x)\} = \Sigma b_{1j} \frac{\partial}{\partial x_j} \{ \Sigma g_{i2}(a_i \cdot x)\}$$

or

$$\text{(2.4)} \qquad \sum_{i=2}^{n} (c_i \cdot a_i) g_{i1}'(a_i \cdot x) = \sum_{i=2}^{n} (b_i \cdot a_i) g_{i2}'(a_i \cdot x) .$$

Let G_i be any function of one variable such that $G_i' = g_{i1}$. Then

$$\text{(2.5)} \qquad \begin{aligned} &\left\{\beta \Sigma b_{1j} \frac{\partial}{\partial x_j} + \gamma \Sigma c_{1j} \frac{\partial}{\partial x_j}\right\}\left\{ \sum_{i=2}^{n} G_i(a_i \cdot x)\right\} \\ &= \sum_{i=2}^{n} \beta[b_1 \cdot a_i] g_{i1}(a_i \cdot x) + \gamma[c_1 \cdot a_i] g_{i1}(a_i \cdot x) . \end{aligned}$$

Integrating (2.4) gives, for some constant k:

$$\sum_{i=2}^{n} c_1 \cdot a_i\, g_{i1}(a_i \cdot x) = \sum_{i=2}^{n} b_1 \cdot a_i\, g_{i2}(a_i \cdot x) + k\,.$$

Substituting this in (2.5) gives

$$\sum_{i=2}^{n} b_1 \cdot a_i \left[\beta g_{i1}(a_i \cdot x) + \gamma g_{i2}(a_i \cdot x)\right] + \gamma k\,.$$

If $k = 0$, then a particular solution to $P_1 f = \Sigma g_i$ would be $f = \sum_{i=2}^{n} (b_i \cdot a_i)^{-1} G_i$. Note we can assume that b_i is chosen such that $b_i \cdot a_i \neq 0$ for all $i = 2, \ldots, n$. The next job is to show how to modify G_i to take care of non-zero k. Let

$$F_i(x_1, x_2, x_3) = G_i(a_i \cdot x) + \delta_i\, e_i \cdot x$$

where δ_i and $e_i = (e_{i1}, e_{i2}, e_{i3})$ will be chosen later. Then,

$$\left\{\Sigma\beta b_{1j} \frac{\partial}{\partial x_j} + \gamma c_{1j} \frac{\partial}{\partial x_j}\right\}\{\Sigma F_i(x_1, x_2, x_3)\}$$

$$= \sum_{i=2}^{n} (b_{1j} \cdot a_{ij})\left[\beta g_{i1}(a_i \cdot x) + \gamma g_{i2}(a_i \cdot x)\right] + \gamma k + \sum_{i=2}^{n} \delta_i \{\alpha b_1 \cdot e_i + c_1 \cdot e_i\}\,.$$

Set $\delta_i = 0$ for $i \geq 3$. Choose e_2 such that $e_2 \cdot b_1 = 0$ and $e_2 \cdot c_1 \neq 0$. Set $\delta_2 = -k/\Sigma c_{1j} e_{2j}$. Then, the function

$$f(x_1, x_2, x_3) = \sum_{i=2}^{n} \frac{1}{b_i \cdot a_i}\, G_i(a_i \cdot x) + \delta_2 e_2 \cdot x$$

solves the equation $P_1 f = \sum_{i=2}^{n} \beta g_{i1} + \gamma g_{i2}$. Having obtained a particular solution, we proceed to describe a general solution of $P_1 f = g$ by studying the general solution of the homogeneous equation $P_1 f = 0$. But, it is straightforward that a general solution of the homogeneous equation is of the form

$g(a_i \cdot x)$ where $a_i \cdot b_1 = a_i \cdot b_2 = 0$. Thus the general solution of (2.3) is

$$\sum_{i=2}^{n} \frac{1}{b_i \cdot a_i} G_i(a_i \cdot x) + \delta_2 e_2 \cdot x + g(a_i \cdot x) . \tag{2.6}$$

In (2.6) e_2 was chosen so $e_2 \cdot b_1 = 0 \neq e_2 \cdot c_1$. Now we use the hypothesis that three of the vecotrs a_i, say a_1, a_2, a_3 are linearly independent, and express e_2 as a linear combination of these three vectors. It follows that the general solution can be expressed in the form

$$\sum_{i=1}^{n} H_i(a_i \cdot x) \quad \text{as required.} \quad \blacksquare$$

References

Ash, J. M. (1976). Studies in Harmonic Analysis, Amer. Math. Soc., Providence, RI.

Dotson, W. G. (1968). Decomposability of positive functions on $\mathbb{R}^n$, Amer. Math. Monthly, 350-357.

Friedman, J. and Tukey, J. (1974). A projection pursuit algorithm for exploratory data anlaysis, IEEE Transactions Computers C-23, 881-889.

Friedman, J. and Stuetzle, W. (1981 a). Projection pursuit regression. To appear, Jour. Amer. Statist. Assoc.

Friedman, J. and Stuetzle, W. (1981 b). Projection pursuit methods for data analysis, Stanford Linear Accelerator, Technical Report.

Friedman, J. Gross, E., and Stuetzle, W. (1981). Multi-dimensional spline-like fits. To appear, SIAM Jour. Sci. and Statist. Comp.

Goutschi, W. A. (1963). On inverses of Vandermonde and confluent Vandermonde matrices, Numer. Math. 5, 425-430.

Henrici, P. (1978). Applied and Computational Complex Analysis, Vol. I, Wiley, New York.

Huber, P. (1981 a). Density estimation and projection pursuit methods, Technical Report PJH-7, Department of Statistics, Harvard University.

Huber, P. (1981 b). Projection pursuit, Technical Report, Department of Statistics, Harvard University.

Lorentz, G. G. (1966). Approximation of Functions, Holt, Rinehart, and Winston, New York.

Lorentz, G. G. (1976). The 13th problem of Hilbert, in, Mathematical Developments Arising from Hilbert Problems, Proc. Symp. Pine Math. XXVIII, Amer. Math. Soc., Providence, RI.

Vitushkin, D. G. (1977). On representation of functions by means of superpositions and related topics, L'Enseignement Math., July-Dec. 1977, 256-319.

Zygmund, K. (1959). Trigonometric Series, 2nd ed., Cambridge University Press, Cambridge, England.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER 313	2. GOVT ACCESSION NO. AD-A115 347	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A NOTE ON NON-LINEAR FUNCTIONS OF LINEAR COMBINATIONS		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL REPORT
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Persi Diaconis and Mehrdad Shashahani		8. CONTRACT OR GRANT NUMBER(s) N00014-76-C-0475
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Statistics Stanford University Stanford, CA 94305		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS NR-042-267
11. CONTROLLING OFFICE NAME AND ADDRESS Office Of Naval Research Statistics & Probability Program Code 436 Arlington, VA 22217		12. REPORT DATE FEBRUARY 2, 1982
		13. NUMBER OF PAGES 16
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

Also distributed under National Science Foundation Grant MCS80-24649, Technical Report No. 179, November 1981.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Projection pursuit; Non-linear functions; Approximation theory.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

PLEASE SEE REVERSE SIDE.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-LF-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

#313

A NOTE ON NON-LINEAR FUNCTIONS OF LINEAR COMBINATIONS

Projection pursuit algorithms approximate a function of p variables by a sum of non-linear functions of linear combinations;

$$(1) \qquad f(x_1,\ldots,x_p) \doteq \sum_{i=1}^{n} g_i(a_{i1}x_1 + \ldots + a_{ip}x_p) .$$

We develop some approximation theory, give a necessary and sufficient condition for equality in (1), and discuss non-uniqueness of the representation.

DATE
FILMED

7-8